

Laboratory Voice Data Entry System

Jeremy L. Praissman and John C. Sutherland

Brookhaven National Laboratory, Upton, NY, USA

Reprinted from BioTechniques 27:1202-1208 (December 1999)

ABSTRACT

We have assembled a system using a personal computer workstation equipped with standard office software, an audio system, speech recognition software and an inexpensive radio-based wireless microphone that permits laboratory workers to enter or modify data while performing other work. Speech recognition permits users to enter data while their hands are holding equipment or they are otherwise unable to operate a keyboard. The wireless microphone allows unencumbered movement around the laboratory without a "tether" that might interfere with equipment or experimental procedures. To evaluate the potential of voice data entry in a laboratory environment, we developed a prototype relational database that records the disposal of radionuclides and/or hazardous chemicals. Current regulations in our laboratory require that each such item being discarded must be inventoried and documents must be prepared that summarize the contents of each container used for disposal. Using voice commands, the user enters items into the database as each is discarded. Subsequently, the program prepares the required documentation.

INTRODUCTION

The storage and analysis of digital experimental data acquired by computer-controlled instruments is playing an ever increasing role in scientific laboratories. Computer database programs are also being used to organize the information needed to run a laboratory (1). In addition, experimental protocols and other documents that are generated by people, rather than instruments, are being prepared in digital form with increasing frequency, typically using a word processor or a spreadsheet program.

One area that has been relatively slow to convert to a digital format is the preparation or modification of documents in which data entry or annotation must be performed during the course of some other activity, such as the preparation of samples. It occurred to us that recently introduced software that permits computers to recognize and act on human speech might facilitate the greater use of digital data during activities in which

the use of a keyboard or pointing device is awkward or distracting. Thus, we assembled a computer system to perform speech recognition at the same time the user moves freely around the laboratory. We also developed a prototype database to perform functions while workers enter data without shifting their hands from experimental samples or apparatus to traditional digital input devices. Here, we describe the system for speech recognition that we assembled and the database application that we wrote to take advantage of voice data entry.

MATERIALS AND METHODS

Computer and Audio Systems

The components required for the voice data entry system are shown in Figure 1. We used a generic personal computer with a 300 MHz Intel (Santa Clara, CA, USA) Celeron processor, 128 Mb of random access memory (RAM), the Microsoft (Redmond, WA, USA) NT operating system (version 4, service pack 4) equipped with Microsoft Office 97 Professional Edition; Dragon NaturallySpeaking® Preferred version 3.52 software (Dragon Systems, Newton, MA, USA) and a Sound Blaster Pro audio interface system (Creative Labs, Milpitas, CA, USA). The microphone headset for input to the audio system supplied with the voice recognition software was replaced by the base unit of an FM Wireless Microphone System (Tandy/Radio Shack, Fort Worth, TX, USA). This system operates in the 6 m fixed/mobile communications band. The microphone and battery-powered transmitter of the wireless microphone system are worn by the person using the data entry system. The transmitter, which measures 60 × 90 × 18 mm and weighs 130 g, can be kept in the pocket of a laboratory coat or attached to the belt. The microphone is connected to the transmitter by a 700 mm long lightweight insulated wire cable, which also serves as the unit's antenna. When attached to the collar of a shirt or blouse, the microphone is positioned close to the user's throat. In some experiments, we replaced this microphone and transmitter with a radio transceiver (21-410 or 21-407; Tandy/Radio Shack) that operates on one of five predetermined frequencies in the same band as the wireless microphone system. The 21-407 transceiver is equipped with a headset that positions a small headphone over one ear and the microphone on a boom near the wearer's mouth. In contrast, the 21-410 combines the functions of a microphone and speaker in a single transducer that is mounted in an earplug worn in the user's ear. Unlike the transmitter that is supplied with the wireless microphone system and broadcasts continuously, both transceivers broadcast only on command in either a "push-to-talk" or a "voice-activated" mode. They can easily be modified to transmit continuously, albeit with reduction in battery life. Both transceivers connect to their audio transducers and attach to the user's clothing in a manner similar to the supplied transmitter, but are slightly larger and heavier.

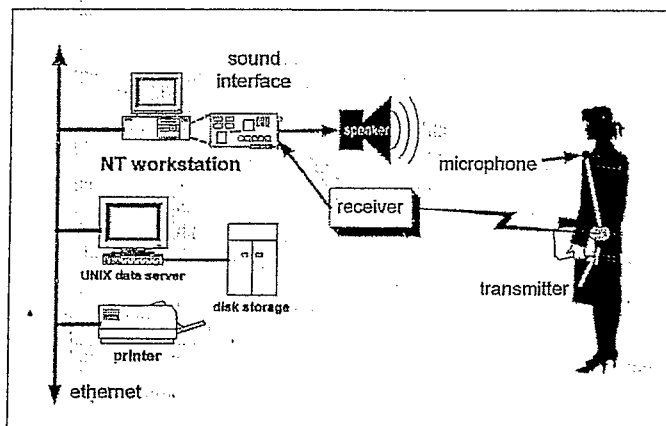


Figure 1. Diagram of the voice-controlled data entry system.

The personal computer used in this application is connected to a local-area network (LAN) and functions as part of a system for data acquisition, storage and analysis, which has been described elsewhere (2). Thus, data files can be stored on a server and later recalled by any person with proper authorization and access to the LAN or the Internet. Various components attached to the network, such as laser printers, are used by the system described here for the generation of reports.

Radiological Waste Inventory Data System

To demonstrate the use of voice recognition in a laboratory environment, we developed a system to track radiological wastes that were discarded during the preparation of experimental samples. We implemented the waste inventory system as a relational database using Microsoft Access 97, which is a component of Microsoft Office 97 Professional Edition. The scheme of the database is shown in Figure 2. The central table of this database records transactions in which a worker discards one or more items of a particular type into a particular container. The other tables in the database contain supporting information about the properties of the objects that can be discarded, the people performing the action, the containers in use, the types of containers available, the types of materials that can be placed in different types of containers, etc.

The various fields available in each table are also shown in Figure 2, as are the dependencies between tables. The transaction form for entering the information required to fill out one row in a Radioactive Waste Inventory Sheet (RWIS) is shown in Figure 3. (In the lexicon of Access and other database management systems, a form is a collection of information formatted primarily for display on the computer screen, while a report is a collection of information formatted primarily for printing. Because the paper-based waste dis-

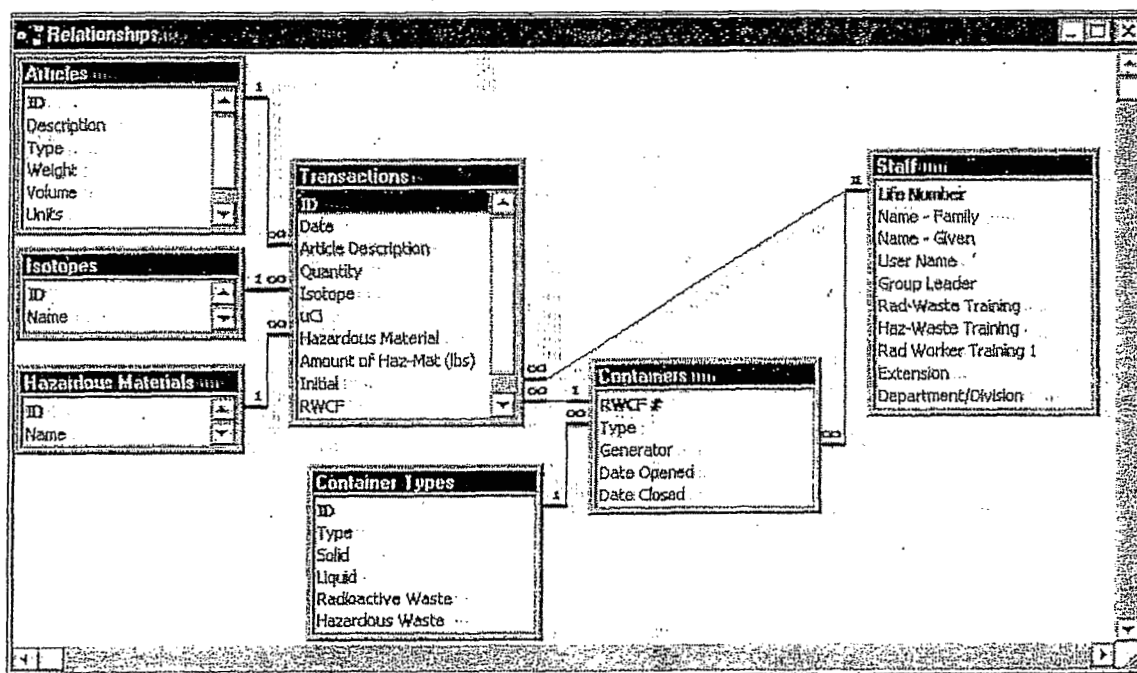


Figure 2. Scheme of the waste inventory database. Data are stored in 7 tables, which are related to one another as indicated by the lines. The names of the tables are indicated at the top with the list of fields given below. The relations are all one-to-many, indicated by the symbols 1 and ∞, respectively. For example, each staff member may be responsible for several containers. Five of the tables hold information used to provide the options available for completing each field of a single transaction in which one or more items are discarded (i.e., a single row of the transactions table).

posal system of our laboratory does not make this distinction; the piece of paper called a Radioactive Waste Control Form [RWCF] is generated as a report in the data system.)

Large fonts are used so that operators can confirm entries easily at some distance from the computer display. The operator can use voice commands to move the focus of the computer screen to various objects (controls) on the form, select items from pick lists that some controls contain or enter data when the range of acceptable values are not known a priori, and to execute other commands. All such functions can also be performed using a keyboard or pointing device, which are provided. Other forms are used when containers are logically opened and closed and to perform other functions.

Table 1 lists the system's forms and reports. Logically, closing a container automatically generates a report that provides a summary of its contents and contains the information that must accompany it through the waste management system. Figure 12 shows a portion of such a report, which is formatted to duplicate the paper RWCF.

DISCUSSION

Our experiments demonstrate that voice data entry in a laboratory environment is possible using readily available hardware and software and offers significant advantages for certain activities. We have also identified factors that should be considered in evaluating the use of voice entry of laboratory data.

Speech Recognition System

Within the past few years, several vendors have introduced modestly priced voice recognition software. Voice entry also requires audio input hardware (i.e., sound cards) that are being supplied as standard components with many personal computers. Likewise, the radio system that permits free movement around the laboratory is designed for the consumer market and is thus reasonably priced. Paradoxically, the most significant cost in equipping a computer system for voice data entry is not in any specialized hardware or software but in the computer itself. Speech recognition is a computationally

RWCF 35996	
Package Type	Mixed-LLRW Plastic Bag, 6ft ³
Enter Natural Text	5 15 ml tubes
Article Description	15ml tube
Quantity	5
Units	Each
Isotope Name	None
Hazardous Material	None
Amount of Haz. Mat	0 LBS.
Parse	OK
Select container	View RWCF

Figure 3. The form used to enter individual transactions into the waste inventory database. The name of the person using the database is supplied to the program from information acquired by the operating system when the person logged in. The type of items that can be discarded are provided in pick lists. Other tables in the database (Figure 8) record information about the discardable items, such as the units in which they are measured, their average unit weight and class (paper, plastic, metal, etc.). Such information is required to provide summaries of the contents of filled containers, as required by existing regulations. The date and time of each transaction are entered automatically when the transaction is completed using data from the computer's clock. The user supplied data (Article Description, Quantity, etc.) can be entered by moving from one control to the next. However, the default data entry method is for the user to speak a phrase such as "numeral 5 fifteen mL tubes". After the voice recognition software interprets these words, the phrase appears in the Enter Natural Text control. If the phrase was recognized correctly, the user activates the Parse control (Parse is the default entry on this form; so one option is to say "new line"), thus activating a code module that parses the phrase and loads the results into the relevant text controls. Both the recognized phrase and the resulting text control entries are shown in the figure above. After checking the entries, the user completes the transaction by activating the OK control (one option is to say "type alt K").

Laboratory Voice Data Entry System

Table 1. Forms and Reports Used in the Waste Inventory Data System

Name	Type	Description
Main Menu	form	Allows selection of the other forms and termination of the program.
Select Container	form	Allows the users to select the container into which items will be discarded from the list of presently open containers. Also provides for the logical closing of a filled container, or opening the New Container Form.
New Container	form	Allows a user to logically open a new container entry in the database. A number corresponding to an Radiological Waste Control Form (RWCF) identifies each such container.
Transaction	form	Contains controls for entering data describing an item or items being discarded. The information from this form constitutes one record in a Radioactive Waste Inventory Form.
Add Article	form	Adds a record describing a new class of items that can be discarded into the Articles table.
Close Check	form	Confirms that the user has specified that a particular container is to be logically closed.
Generate Report	form	Initiates generation of a Radioactive Waste Inventory report or an RWCF, which are described below.
Radioactive Waste Inventory	report	Lists all of the records of items placed in a container plus header information such as the type of container and the name of the person who generated the waste. A portion of a Radioactive Waste Inventory prepared by the Waste Inventory Data System is shown in Figure 5.
RWCF	report	Summarized the contents of each container. An example is shown in Figure 6.
Articles	report	Lists the contents of the Articles table (i.e., the classes of articles that can be entered into the database as having been discarded). This report is generated by a control located on the Add Articles Form.

demanding process and requires both a fast processor and a lot of RAM. The system we assembled frequently requires a few seconds to recognize some words and phrases, a delay that can be distracting during an experimental procedure. Thus, we recommend that voice data entry systems should be based on the fastest available microprocessor. Some recently introduced microprocessors have special on-chip hardware that facilitates computationally intensive multimedia tasks. Presumably, such systems would be worth the additional cost in voice data entry applications.

The microphones we tested represent a range of trade offs. The headset supplied with the voice recognition software appeared to provide the highest quality sound and would be adequate in applications in which the operator remains seated in one location (e.g., while operating a microscope). However, the connections to the computer are too restrictive if the operator must move around the laboratory. The clip-on microphone supplied with the FM wireless microphone system is unobtrusive, but when attached to a person's collar, it is positioned near the wearer's throat rather than near the side of the mouth, as recommended by the speech recognition software documentation. The FM radio transceiver with a headset positions the microphone properly, but is intrusive. The transceiver with the microphone in an earplug provides adequate audio quality with minimum bulk and reduced the effects of room noise. However, some users found the earplug uncomfortable. All of these transmitters provided signals that the speech recognition software was able to process correctly. The software documentation recommends retraining the program whenever the microphone or other components of the audio system are changed.

The accuracy of speech recognition software is influenced by background noise, thus a quiet environment is desirable. We should have liked to use the text-to-speech capability of the software to provide feedback to the user indicating the data was entered, but this capability was not available for use with Access in our version of the software. However, we use large fonts in the transaction form to permit operators to check the data even when they are located some distance from the computer. We also programmed the computer to generate a bell-like sound to confirm that a record has been entered. Text-to-speech using the speakers of the sound system would prevent more than one person from using voice entry in a given area. However, multiple users could be accommodated, if the speaker system was replaced by a radio link from the computer to the personal headphones of the user. Our experience with the radio transceivers suggests that using the same frequency for both directions would not be desirable. Pressing the "push-to-talk" button of the transceiver defeats the goal of hands-free data entry, while the "voice-operated-transmission"

Figure 4. Add Articles form used to add a new class of discardable items to the data base. This form allows end users to add new classes of articles to the Articles table (Figure 2), which provides the pick list for the Article Description control on the Transaction form (Figure 3). Records can also be added to or removed from the Articles table and the five other supporting tables shown in Figure 2 by opening the table in Datasheet View using the Access programming environment, but such procedures are normally performed only by the database manager.

mode causes the user to preface most phrases with "aggggah" or some similar sound to initiate a transmission, thereby confusing the speech recognition program. Thus, user-specific voice feedback would be most effective if separate radio frequencies were used in each direction, i.e., full-duplex mode.

Speech recognition is a statistical process, so recognition errors at some level are unavoidable. Effective operation of the speech recognition software requires significant training of both the operator and the speech recognition program. The speech recognition program is "trained" by having each user read text displayed by the program on the computer screen. The system maintains a separate voice recognition resource file for each user. There is also a provision for building application-specific vocabularies by providing files that contain technical

terms and by responding to questions from the program regarding words not in its dictionary. Such training is also specific to each user.

Voice recognition is, in some ways, like a "command line" computer interface, in that the operator needs to know about a dozen spoken commands that are used to control the computer. These requirements mean that voice data entry is best suited for applications in which a small number of people frequently perform repetitive tasks that benefit significantly from capturing data in digital form as the work proceeds. An example of such a program is described below.

Radioactive Waste Inventory

Package type: Mixed LLRW Packaging, 6072 RWCF #: 15096

RWCF contents summary:

Date	Article Description	Quantity	Isotope Name	uCi	Hazardous Material	Amount of Haz-Mat	Initial
10/09/92 14:13 PM	Canister, metal	3	Uranium	0	None	0	Low
10/09/92 02:40 PM	Canister, metal	1	U-235	2	None	0	Low
10/09/92 02:46 PM	Canister, metal	103	Uranium	0	Azodic	0.1	Low
10/09/92 03:03 PM	Canister, metal	7	U-235	1	None	0	Low
10/09/92 03:03 PM	Canister, metal	3	Uranium	0	None	0	Low
10/09/92 03:03 PM	Canister, metal	2	Uranium	0	None	0	Low
10/09/92 03:03 PM	Canister, metal	12	Uranium	0	None	0	Low
10/09/92 03:03 PM	Canister, metal	10	U-235	2	None	0	Low
10/09/92 03:03 PM	Canister, metal	6	Uranium	0	None	0	Low
10/09/92 03:03 PM	Canister, metal	1	Uranium	0	None	0	Low
10/09/92 03:03 PM	Canister, metal	2	Uranium	0	None	0	Low
10/09/92 03:03 PM	Canister, metal	1	U-235	3	None	0	Low
10/09/92 03:03 PM	Canister, metal	103	Uranium	0	None	0	Low
10/09/92 03:03 PM	Canister, metal	103	U-235	12	None	0	Low
10/09/92 03:03 PM	Canister, metal	103	Uranium	0	None	0	Low
10/09/92 03:03 PM	Canister, metal	103	U-235	12	None	0	Low

Figure 5. Radioactive Waste Inventory. This report is generated when the container designated by the indicated RWCF is logically closed. It lists all of the items placed in the container including data automatically added by the computer system such as the date and time of entry and the person recording the transaction, plus "header" information such as the type of the container and the RWCF number. It can be printed or displayed on a computer screen, as shown here.

Radioactive Waste Control Form RWCF #: 15096

Generator: John Doe

Department/Division: Biology Ext. 2194 Inlet/Outlet #: 1200

Dept. Representative for Waste: Biology Sect. # At work Area B Bldg. # 451 Placement Date: 9/2/99

Package Type: Mixed LLRW Packaging, 6072 Return Package? YES ☐ NO ☒

Volume of Waste: Liquid (g/L) OR Solid (kg) 0.020 ft³

Describe Waste:

Physical State: Solid ☒ Liquid ☐ Gas ☐ Is waste eligible for Decay-In-Storage? (30 days) ☐ yes ☐ no

Components and their percentages by volume:

WEIGHT OF WASTE: 336 kg	
Paper	Percentage By Volume: 41%
Sludge	Percentage By Volume: 39%

Radioactive Isotopes and activity of each in microcuries:

U	235	uCi
---	-----	-----

Analysis method used to determine activity: Direct Method ☒ Indirect Method ☐

Canister closed after periodic documentation: Gamma Ray Spectroscopy ☐ Direct Read to Activity Conversion ☐

Figure 6. Part of a report generated by the waste inventory database system summarizing the contents of a specific container. The summary information is generated automatically, saving the user the time and effort required to compile such information by hand. It can be printed or displayed on the screen, as shown here.

Laboratory Applications of Speech Recognition

Our prototype Waste Inventory Data System database application has many of the attributes appropriate for voice data entry in a laboratory environment. Experiments involving the handling of radionuclides take place in reserved areas that can be entered only by authorized individuals, hence reducing interference from background noise. Workers handling radionuclides must complete several types of training, so the additional invest-

ment in learning the use of the voice data entry system is a small fraction of the total training load. The requirement that items discarded during such experiments must be inventoried, whether or not they are contaminated, places a premium on capturing such data as the items are discarded. The types of items discarded during an experiment are usually limited (tubes, paper wipes, pipet tips, etc.) and thus can be easily entered in lists from which the user can make a selection. New items can be added to the list using the "Add Articles" form, which also records information on the items' weight, type (paper, plastic, metal, etc.), state (solid or liquid) and the name that will be used in voice recognition, as shown in Figure 4. The major advantage of the voice data entry system is realized when a container is "closed" using the Select Container Form. This step calls a code module (written in Access Visual Basic for Applications) that moves the entry for each item in the transactions table to the Radioactive Waste Inventory report, as shown in Figure 5. The information in this report is also condensed to produce an RWCF, which summarizes the contents of the container. Figure 6 shows an example of this report, which can be printed and attached to the container when it is delivered to the waste management system. The code module performs a variety of required calculations, such as determining the fraction of the container composed of paper, plastics, metals, etc.; hence eliminating considerable time that a person would take to perform these calculations.

An advantage of implementing the radiological inventory system using a relational database as opposed to, for example, a spreadsheet, is the ease of checking that organizational rules are followed. For example, entering a transaction in which a liquid sample is reported as being placed in a type of container not designed for liquids generates a warning message. We also check that the person generating the waste is currently certified to handle the materials being discarded. This check is implemented using the various training fields in the Staff Table (see Figure 2), which contain the date through which a staff member is authorized to perform a given function. At present, such data must be maintained by the local database administrator, but could be obtained by our database program by submitting a query over the network to the laboratory-wide training database. Similarly, the data listing the contents of each container and the various summaries contained in the RWIS and RWCF, respectively, could be transmitted directly in digital form to a central database.

CONCLUSION

Speech recognition software can be used to enter data during an experiment without diverting the user from the task at hand. Replacing the microphone supplied with speech recognition programs permits the user to move freely around the laboratory. While the cost of the auxiliary hardware and software to implement such a system is modest, the computationally demanding nature of speech recognition requires a fast microprocessor and a large quantity of RAM to provide acceptable performance. Effective use of speech recognition in a laboratory is facilitated by a quiet environment and requires significant training of both the people using the system and the speech recognition software. Thus, speech recognition is most useful for applications in which real-time capture of data during a procedure provides substantial benefits. The recording of items discarded during work involving radioactive or hazardous samples is such an application because the database program, rather than a staff member, prepares the documentation required by waste management regulations.

ACKNOWLEDGMENTS

We thank Denise C. Monteleone and John G. Trunk for assistance in configuring the computer system, Paula V. Bennett and Robert L. Colichio for explaining the operation of the radiological waste protocols and evaluating the operation of our prototype radioactive waste database and Ann Emrick for assistance in preparation of the manuscript. J.L.P., a computer science major at Carnegie-Mellon University, was the recipient of an Energy Research Undergraduate Laboratory Fellowship from the United States Department of Energy. Our research was supported by the Office of Biological and Environmental Research, Office of Science, United States Department of Energy.

REFERENCES

1. Naylor, O. and S. Stamm. 1999. ScienceLabDatabase: a computer program to organize a molecular biology laboratory. *BioTechniques* 26:1186-1191.
2. Sutherland, J.C., D.C. Monteleone and B.M. Sutherland. 1997. Computer network for data acquisition, storage, and analysis. *J. Photochem. and Photobiol. B* 40:14-22.

Update to: Laboratory Voice Data Entry System

John C. Sutherland

*Brookhaven National Laboratory, Upton, NY and East Carolina University,
Greenville, NC, USA (sutherlandj@mail.ecu.edu)*

Voice recognition, combined with database technologies, offers the potential of increased efficiency in the capture and reporting of a variety of data in the laboratory environment. It is particularly valuable in situations in which repetitive and, hence, predictable operations must be performed and where ordered and predictable data must be stored for future reference. The gains in productivity that can be achieved in such environments can justify the costs associated with developing voice recognition database systems. In appropriate circumstances, increases in the ease of data entry outweigh the inherent loss in flexibility that such systems impose.

The technologies underlying voice data entry systems has improved significantly in the short time since we described this system. Voice recognition technology, the most application-specific enabling technology, has shown significant improvement with several vendors offering excellent systems. Dragon NaturallySpeaking® voice recognition software is currently shipped as version 6, compared to the version 3.5 product that we used. Among the most significant improvements are the ability of the current version to correctly recognize words spoken as a part of a continuous stream, compared to the "___ distinct ___ word ___ pattern ___ required ___ in ___ nineteen ___ ninety ___ nine." The available tools for developing voice recognition applications have also improved.

There have been important advancements in several other supporting technologies. Computer performance has forged ahead, as we have come to expect. The rapid penetration of mobile phones in the communications marketplace has brought with them a wide assortment of improved audio input/output gear—headphones and microphones—which are both more comfortable and offer improved fidelity at modest cost. Computer displays represent a less obvious, but perhaps more important, area of improvement. The cost of flat screen monitors has dropped sharply in the last few years, and these devices are far less obtrusive and safer in the laboratory workspace than their cathode-ray tube-based predecessors.

Despite these significant improvements, voice recognition is still only a niche form of human-computer interface. More training and practice are required to enable effective use of voice recognition input compared to the ubiquitous keyboard and mouse. Thus, voice recognition is still most likely to be successful in situations in which laboratory personnel regularly perform generally similar operations in reasonably isolated and quiet areas.

Voice data entry will be most likely to return benefits greater than cost when it is part of a comprehensive systems approach to facilitating laboratory operations, rather than an independent addition to existing procedures. Adding voice data entry to existing procedures is particularly problematic if the organizational entity developing the new capability is not in complete control of the entire process. For example, developing a database to fill out one or more forms or reports should not be attempted unless the authority to specify the format of the report rests with the same entity that develops the data system. Otherwise, the effort expended in implementing the data system may be nullified by changes in the underlying process or the reporting requirement.